PB 198 515

A Comparative Analysis of the Relationship between Mortality and Air Pollution -Weather

FINAL CONTRACT REPORT

by James R. McCarroll, M.D.
Principal Investigator

Michael D. Lebowitz, Ph.C., Research Associate
Department of Environmental Health
School of Public Health and Community Medicine
University of Washington
Seattle, Washington 98105

I. Utilizing Multiple Regression Analysis

Abstract:
Multiple regression analysis, with lags, was performed on daily mortality on air pollution and weather variables for three United States cities (New York, Philadelphia, Los Angeles). Relationships between mortality and air pollution-weather were found for all three cities. These relationships were consistent within cities, by season. They were also consistent between cities, allowing for generalization about the relationship: daily mortality is associated with air pollution and weather in urban areas of the U.S. This association can be stated as a formal relationship in time. The effect of pollution and weather on mortality is essentially an immediate one.

This research was supported under Contract No. CPA-22-69-126, National Air Pollution Control Administration, PHS.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

DISCLAIMER

This report was furnished to the Air Pollution Control Office by
Department of Environmental Health
School of Public Health and Community Medicine
University of Washington
Seattle, Washington 98105
in fulfillment of Contract No. CPA-22-69-126

PB 198 516

BIBLIOGRAPHIC DATA SHEET	1. Report No. APTD-0650	2.	3. Recipient's Accession No.
4. Title and Subtitle A Comparative Analysis of the Relationship Between Mortality and Air Pollution and Weather -- I. Utilizing Multiple Regression Analysis.			5. Report Date June 1965 6.
7. Author(s) James McCarroll, & Micheal D. Lebowitz			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Department of Environmental Health School of Public Health and Community Medicine University of Washington Seattle, Washington 98105			10. Project/Task/Work Unit No. 11. Contract/Grant No. CPA-22-69-126
12. Sponsoring Organization Name and Address EPA, Air Pollution Control Office Technical Center, Box 12055 Research Triangle Park, N. C. 27709			13. Type of Report & Period Covered 14.

15. Supplementary Notes

16. Abstracts

Multiple regression analysis, with lags, was performed on daily mortality on air pollution and weather variables for three United States cities (New York, Philadelphia, Los Angeles). Relationships between mortality and air pollution weather were found for all three cities. These relationships were consistent within cities, by season. They were also consistent between cities, allowing for generalization about the relationship: daily mortality is associated with air pollution and weather in urban areas of the U. S. This association can be stated as a formal relationship in time. The effect of pollution and weather on mortality is essentially an immediate one.

17. Key Words and Document Analysis. 17a. Descriptors

Air pollution
Urban areas
Analysis (mathematics)
Weather

17b. Identifiers/Open-Ended Terms

Mortality
Multiple regression Analysis

17c. COSATI Field/Group 13/B

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Unlimited	UNCLASSIFIED	26
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-38 (10-70)

USCOMM-DC 40329-P71

The major episodes of Donora (1), London (2), Meuse Valley (3), and others, were the first large attempts at air pollution epidemiology (4, 5). Research on the relation between mortality and environmental exposure have been characterized by three different approaches, which differ in sophistication as well as methodology.

The spatial approach is one in which comparisons are made between mortality levels in areas characterized by different pollution levels. These areas may be parts of regions or countries or may be sub-areas of a city. Comparisons between larger areas are represented by Daley (6) in England, by Shiffman and Landau (7) and by Hickey, et al. (8) in the U.S. Within city comparisons have been made by Winkelstein (9,10) in Buffalo, Ziedberg (11) in Nashville, Lepper, et al. (12) and Carnow (13) in Chicago, and Rumford (14) in Philadelphia and Chicago.

As Rumford has stated: "The intra-city and intra-area investigations are usually based on the assumption that if there is an association between manifestations of air pollution and mortality, then residents living in areas of high pollution should experience higher mortality rates. Because polluted areas usually have other adverse environmental conditions present, the total social-economic complex of the area must be considered." Such considerations are crude, though they are found in the work of some authors (Winkelstein, Zeidberg, Lepper). In some of these studies, more advanced statistical methods are used, such as multivariate regression (Hickey). However, as seen above, more of the studies use long-term averages of mortality and pollution rather than continuous, multivariate approaches, and they ignore many sources of internal variation.

The studies of temporal variation go somewhat beyond the above in their use of more advanced techniques and detailed temporal data to look for more detailed relationships between air pollution and mortality. Thus Scott, et al. (15), in London looked at daily winter mortality and correlated it with temperature, SO_2 and smoke variations, using lags and partial correlation coefficients. Boyd's work (16) in England is also well-known, in which he went further than Scott by looking at spatial variations also, as well as examining age and cause-of-death specifics.

A statistically more advanced type of this study was performed by Hechter and Goldsmith (17) in Los Angeles, utilizing Fourier Series and Spectral Time Series.

By using time series, the authors could study auto-correlation effects of residuals and lags. They removed the seasonal component as well as trend before relating the mortality to the temperature and pollution. Using continuous distributions, they examined total mortality, cardiac and respiratory deaths for Los Angeles from 1956 to 1958, and compared this to oxident, temperature, and carbon monoxide from the environment. Their analysis did not find any association between changes in daily air pollution level as an influence upon day to day mortality patterns. It is worthwhile to note some of the major points they brought out in this paper, such as: "Mortality is at best a rough indicator of an acute response. Death is a terminal event which follows a long sequence of circumstances; its occurrence at a particular time is attributed in large part to a person's general disease history. The lack of an easily measurable and sensitive biological indicator is a serious limitation to ascertaining the effect from low level air pollution. Mortality is commonly used as an indicator because it is easily measurable, but it is far from sensitive." They also point out the value of looking at lags, since "when the variables are studied over time, the most meaningful association is not necessarily obtained by correlating the simultaneous values of the variables; there could be a delay in the influence of one variable to another."

Martin and Bradley (18) in London, and later Greenburg (19,20) in New York, specifically looked for excess mortality during episodes of air pollution (with or without extreme weather conditions). Both considered cause-of-death and the latter looked at age differences. However, their methods were of the discrete, comparisons type, and did not examine more subtle relationships over time.

Both the Martin and Bradley and the Greenburg, et al. papers discuss specific excess mortality at specific times utilizing deaths by day, the meteorological observations for the areas at that time, and the atmospheric pollutants measured, in a post-hoc study of excess mortality. They conclude that atmospheric conditions have an effect in determining the time at which the mortality occurs and the time at which it may reach its peak (cf. Martin and Bradley). This emphasizes again the need to find a way to study the role of pollutants in the pathogenesis of disease.

None of these earlier studies of excess mortality, however, gave much attention to the effect of place-of-residence or to the long term environmental conditions associated with mortality.

Glasser and Greenburg (21) studied total mortality by day, for 1960-64, in relationship to SO_2, smoke shade (PM), temperature, wind speed, and sky cover. They used 15-day moving averages and deviations from normal to transform the data. They used regression analysis as well as cross tabulations. They conclude that "SO_2 was found to be positively associated with mortality independent of the weather factors studied. The evidence suggests that this increase in mortality was associated with an increase of SO_2 at all levels [with regression, if increase is linear, association is linear, i.e., at all levels]." Relationship was also found with smoke shade, temperature and occasionally wind speed and sky cover (none was related to mortality for all analyses).

Hodgson (22) used New York City data for the three years, 1962-65, to relate monthly average mortality to degree days (a deviation figure), SO_2 and particulate matter, for mortality ages 65+ as well as under 65, using multiple regression analysis; he also separately analyzed specific causes-of-death. He did not control in any way for season or day-of-week, etc., nor did he use other meteorological variables. He did not measure the effect of using monthly averages over the years, which may be a large factor since all the variables are cyclic over months and have similar cycles. He did use moving averages, which as he says, prevent precise estimation, and his use of simple multiple regressions does not control for the interactions (high correlations) present in his independent set of variables. He looks at lag effect but does so by month, which must be extremely insensitive since these variables are all quite variable by day. He discusses the cumulative effect of exposure which, while quite interesting in regard to my work, is also highly insensitive for the same reasons. There are only 36 points in his monthly regressions (the number of months in three years) which also inflates his results. The use of degree days based on an average temperature of 65 for the whole year also embodies unreasonable assumptions about variability (i.e., he forgets the seasonal effects, etc.). His regressions on daily mortality were for all seasons, with no controls for extraneous variables and with only a few independent variables. The results were positive but the degree of association minimal. He concludes that temperature, particulate matter and SO_2 are all related to mortality (especially the first two for the 65+ mortality and the second for the under 65 mortality). He concludes, from his regression results,"...there were more deaths, the greater the concentration of atmospheric pollution." On the basis of the size of the relationship (an R of about .6-.7), again related to his small n, etc., he says: "If air pollution causes [sic] the observed variation in mortality, and if particulate matter and SO_2 at least reflect the presence of the harmful environmental

factors, then slight or moderate increases in concentration of air pollution during a month can be expected to result in increased mortality...amounting to several hundred deaths [per month, again based on regression results and the linearity of a relationship therein]." Improvements in the type of regression used and in multi-variation of environmental conditions, as well as greater detail in mortality, and using continuous distribution forms of analysis would be beneficial, in this case, beyond what Hodgson conducted.

McCarroll and Bradley (23), in their study of excess mortality in relation to air pollution and weather extremes in New York, looked for patterns in episodes and for regularly recurring episodes which normally go unrecognized. Their methods were still those of discrete comparisons, though they did look at age differences and looked at the lag effects of both the pollutants and the weather. They found patterns in many of the episodes which were remarkably consistent, and stressed, i) the immediacy of the rise in mortality following a rise in pollution; and ii) the fact that the effect is not confined to the older age group. "These characteristics suggest that if these excess deaths are related to air pollution, the mechanism is probably protean and pervasive...." These last three studies utilized the same or similar data for New York City as this study, and a comparison of the results is made in the Discussion.

Cassell, et al., (24) have concluded"...that, in part, the lack of progress [in discovering the nature of the relationship between excess mortality and air pollution] stems from a continuing and unbecoming simplistic approach...". They claim that there has almost never been a systematic examination and definition of air pollution episodes compared with mortality over prolonged periods. They also state that the examination of specific episodes has required considerable arbitrary definition to isolate the effect of air pollution.

In their examination of three years data for New York, Cassell, et al, first used a sine curve to remove the seasonal effect and then used confidence intervals to find "unusual events" (occuring 5% of the time or less). Most of the previously reported episodes were confirmed. They also confirmed "that elevated pollution levels were not necessarily associated with increased mortality."

Their conclusion is that the pattern of deaths in New York City is very uniform over time. They stress the complex and multi-variable forces involved in a given short period of time which make it difficult to assign any specific "cause and effect" role, and this combined with the long-term stability of mortality levels can lead to numerous interpretations.

It is apparent from the above that these various studies differ in many respects in how they looked at mortality and its relation to air pollution and weather. In regard to mortality itself, many only looked at either the spatial or temporal mode, looked at total deaths rather than cause specific or age-race-sex specific deaths, often did not look at related variables such as age and socio-economic status, often used averages rather than continuous mortality variation. Methodologies involved all seem to differ in important respects. In addition, the weather and pollution variables differed from one study to another, with the exception that the pollutants often incriminated were sulfur dioxide and particulate matter (25). As highlighted by McCarroll and Bradley, however, rises in mortality often did follow rises in air pollution and the effect was not confined to any specific age group. This can be said to be the summary of existing information from these studies.

Cassell, et al, (24) go so far as to say "we believe that, in part, the lack of progress stems from a continuing and unbecoming simplistic approach, which we have shared, in using mortality statistics to show the effects of an environmental factor on health." In short then, it is obvious that we have

to start looking at various continuous, multivariate approaches to the study of mortality and its relation to environmental disturbances, and these disturbances must be brought out to numerical interpretations that differ from the similar mix of forces resulting in remarkably stable patterns of mortality.

Shortcomings of previous studies point out the need for new methods. Not only should one use the multivariate continuous statistical techniques presently available, but one must look for new methods. The new methods should specifically be able to measure a relationship over time in which as many of the important factors are considered and thru which the extent of the environment and the mortality are related in time. It is emphasized that a rather long period must be looked at for any given place and that comparative studies of such relationships must be done concurrently in different places. Thus, variations in mortality, environment and social environment can be considered and similarities or differences between the cities used to further explain the findings. Using different cities, say two similar and one different (in regard to mortality and environment), one can estimate the external consistency of the findings. One can generalize more if at least the two similar cities present similar results. Similarities or differences in the findings for all three cities will tell us more about the overall relationship between mortality and air pollution-weather.

Purposes of the Study:

This study is a further attempt to find a relationship between daily mortality, air pollution, and weather. New methods were devised for this purpose. The study was comparative; three U.S. cities were studied at the same time using the same methods and similar data from each.

The hypothesis under study was: Can one detect a relationship between mortality and air pollution-weather, using available material on the environment and on mortality? In addition, does this relationship vary by season and city?

Materials & Methods

Data for this study covered three years for New York City (1962-65), two years for Philadelphia (1963-4) and four years for the Los Angeles area (1962-65). New York City Mortality data was provided on cards by the Cornell Family Illness Study (CFIS), which obtained the data originally from the City Health Department. Philadelphia mortality was provided by the National Air Pollution Control Administration from tapes it received from the National Center for Health Statistics. Los Angeles mortality data was furnished by the California State Health Department. Information on the weather variables in all cities came from the U.S. Weather Bureau stations in cities. Air pollution data was provided by: the City and CFIS monitoring stations for New York, the Philadelphia city agency for Philadelphia and the Los Angeles Air Pollution Control Agency for Los Angeles.

The variables for which information was available and were used in this study are:

1. Mortality: by age groups;
2. Air Pollutants (daily measures): Sulphur Dioxide (SO_2), Particulate Matter (PM), (NYC's in COH units, Philadelphia's in TSP and COH units, none for LA), Carbon Monoxide (CO) (not for Philadelphia), Oxides of Nitrogen, (NO_x) (not for Philadelphia).
3. Weather variables (daily measures): Temperature, Relative Humidity, Wind Speed and Radiation (the last two only for NYC and LA).

The pollutant measurements were city averages for LA and downtown readings for the other two. In New York, SO_2, PM and CO were measured in lower Manhattan, and NO_x was measured in upper Manhattan. Measurements of specific pollutants in different cities often are based on different instruments, necessitating some transformation for standardization.

A method was desired to delineate both the long-term patterns and any large deviations from them. Seasonal variations and other factors which might produce extraneous variations would have to be identified and controlled. Time trends were not considered since the period covered was short.

The long-term pattern, including seasonal variation, can be described in terms of cyclic functions or seasonal base levels of mortality. One can then study deviations from this pattern or base (26, 27, 28). In the process, random day-of-week variations should be removed (26) but weekly averages should not be employed as they may mask the peaks of mortality for which one is looking (26). The use of moving averages creates artificial cycles, masks daily values and invalidates day-to-day independence (26). Spectral time series analysis is one of the continuous multivariate techniques available (17, 29). However, it yields results not consistent with those of other methods. It may remove too much of the variation in the data, and it does not permit large-scale multivariate analysis. For these reasons, time series analysis has not been used in this study. Instead, a method of transforming the variables was used, similar to that used in the author's previous study of morbidity and air pollution-weather.

In earlier analyses of these data, various periodicities have been demonstrated, which may have been due to either the periodic nature of the day of death, or the variation introduced into meteorological and air pollution measurements by season of year and day of week (e.g., traffic patterns, industrial activity, etc). To remove these variations from our analyses, aimed at studying the relation between daily mortality and meteorological and pollutant variables, each analysis has been made specifically for each consecutive season. Within each season, the variability attributable to the day of the week was removed by dividing the deviations from the day average by the standard deviation. Thus, for the mortality on

a particular Saturday during the spring of 1963, we have as a basic unit for analysis

$$Z = \frac{x - \bar{x}}{s}$$

where x is the mortality for the Saturday in question, $\bar{x}$ is the average mortality for all Saturdays in the spring of 1963, and s is the standard deviation of these daily mortality numbers. This method is derived from the standard normal transformation.

Variations concomitant with season, day-of-week and year (in their broad meaning) may also exist secondarily to these factors. They are ill-defined time related phenomena. These also are removed by the transformation. Furthermore, the transformation of each original variable furnishes a new variable which is a measure of variations from the base level-pattern. The new variable is in standard units so that all similar variables can be compared directly. Comparisons between cities also are facilitated since the transformation standardized the measurements of the variables. Thus, the differences in the instruments used to measure the same variable (such as SO_2) no longer interferes with comparisons between cities. (I.e., each SO_2 value is now a variation from its original base and its measurement is in standard units).

Six-month seasons, distinguishing the months with the highest mortality from those with the lowest mortality, were used in analysis. The seasons were: winters- October thru March and summers- April thru September. Such long periods were needed to provide sufficient numbers of days for adequate determinations of daily means and standard deviations and for analysis by season. Missing data within each season reduced the number of days available. In this study, missing data was always excluded from the parts of the analysis into which it would enter, thru the use of a method similar to that of Afifi and Elashoff (20).

One of the primary methods of multi-variate statistical analysis of continuous data is multiple regression analysis (31). This method furnishes a key to the relationship between the dependent variable (daily mortality) and the set of independent variables (the air pollution and weather variables). It does so by developing a linear equation (the regression line in multi-dimensional space). This equation can be used to "explain" the relationship for existing data or to "predict" the relationship for similar future data. It also yields the multiple correlation coefficient, which is a good indicator of the strength of the relationship. This method is packaged for computer usage (32).

One of the purposes in utilizing this method is to determine the consistency of any relationship within a city and also between cities (measure of internal and external consistency — reliability and validity). It is an excellent method for determining the differences or similarities between cities. Comparative analysis is rare in health-effects studies and should furnish very useful information. The similarities and differences between the cities in terms of their physical environments should facilitate interpretation of the results. Thus, they should clarify the relationship between mortality and pollution-weather in urban settings.

The analysis presented here has considered only the similar set of variables from each city. Thus, total and age-specific numbers of deaths were used as the dependent variables. The independent set of variables included the daily values (averages) of the pollutants: SO_2, PM, CO, NO_x, and the weather variables: temperature, relative humidity, wind speed, and radiation.

All the findings reported here are based on stepwise multiple regression analyses of the raw data and standardized variables described above.

If the daily mortality were linearly related to the total set of available meteorological and pollutant variables, the stepwise procedure would generate the coefficients a_i in the following linear equation:

$$C = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_8 x_8$$

In this equation, C is the standardized mortality measurement and x_1 to x_8 are the standardized measurements of the eight pollutant and meteorological variables one at a time in order of their ability to "explain" the variability in the daily mortality as measured by the square of the multiple correlation coefficient. The procedure terminates when none of the remaining unentered variables significantly increases the explained portion of the variability in the daily mortality. Thus, in any given analysis of the relation between the health measurement and the environmental variables, the procedure could produce an equation with as many as eight terms on the right hand side. The magnitude and sign of the coefficients in the equation determined by the procedure expresses the "strength" and "sense" of the contribution of that variable to the linear determination of the daily mortality. When standardized variables are used, the coefficients may be compared to one another to obtain relative contributions of each of the variables in the same and different season.

Because longer lags did not serve to increase the associations, only one and three day lags have been used routinely. Because many reports have suggested delayed effects of pollution-weather on mortality, the strength of the relationship between the standardized variables was examined using lags up to 15 days for the winter in New York. Only statistically significant regression results will be presented.

Results

One major finding, in terms of the mortality data, is that most of the deaths and most of the variation in mortality occur in the older age group, specifically 65+. (This is quite important for later analysis.) This is similar to the findings in Japan (33). There is a large amount of variation in most of the environmental variables.

Although the numbers of daily deaths differed, the crude mortality rates of the three cities during the study period were similar. New York had an average of 241 daily deaths during the study period, Philadelphia 77 and Los Angeles 159. The standard deviations of daily deaths in the three cities were all approximately one-tenth of the means, indicating small variation. Deaths of those under 5 were so few in number as to be useless for further analysis. Deaths of those under 45 accounted for about 12-18% of total deaths in the three cities. At least 80% of deaths were in the 65+ age group.

The three cities had different levels of air pollution and different meteorology. New York had higher levels of SO_2 than Philadelphia, and the highest levels of PM. Los Angeles had higher leve of CO than New York, but the levels of NO_x were similar for the two cities. On the meteorological front, New York and Philadelphia were slightly colder than Los Angeles. Philadelphia had a slightly higher level of relative humidity than New York or Los Angeles. New York was windier than Los Angeles. Information on CO, SO_2, NO_x, wind speed and solar radiation was not available for Philadelphia.

From the above information, it is hard to determine whether New York and Philadelphia have similar or different environmental conditions, primarily

due to the lack of information about Philadelphia. It will be assumed in this study, however, that they are similar, especially in contrast to their differences with Los Angeles.

The levels of mortality, pollutants and meteorological variables all differ significantly between winters and summers. This indicates that separate analysis for the two seasons is necessary. The levels by day-of-week also show differences, but no significant trend. The pollutants are generally lower on the weekends. The differences in mortality and weather from one day to another are not large.

Regressions: Both the pollutants and the meteorological variables are important in these regressions. The multiple correlations for the total period are above 0.3 for New York and Los Angeles, and above 0.1 for Philadelphia. (Note: Philadelphia correlation coefficients are always lower than the others, in part because there are fewer variables used.) It appears that the correlation coefficients are not consistently higher in the winter or the summer seasons for New York or Los Angeles. The multiple correlation coefficient ranges from 0.38 to 0.83 in New York City, from 0.19 to 0.36 in Philadelphia, and from 0.26 to 0.54 in Los Angeles (see tables 1-3).

Variables that are most often important in the relationships in New York are: sulfur dioxide, particulate matter, temperature, and wind velocity. In Philadelphia, they are: temperature, humidity and sometimes particulate matter. In Los Angeles, they are: SO_2, NO_x, temperature, relative humidity and radiation. In all three cities, temperature and some pollution are consistently important, and both relative humidity and wind velocity are often important (Tables 1-3).

15

During the winters, temperatures are usually low (5 of 7 seasons) and the pollutants are often high (6 of 7 seasons) in all three cities. The pollutants that usually appear significant are SO_2, particulate matter or NO_x in New York (3 of 3 seasons), and both SO_2 and NO_x in Los Angeles (3 of 3 seasons) and COH in all winters in Philadelphia. In New York, wind velocity is often high (2 of 3 seasons). In Philadelphia, relative humidity was low in the one winter considered.

During the summers, high temperatures are significant about half the time in all three cities. Pollution is often significantly high in all three (6 of 8 seasons): in New York, both summers had high pollution, one of two did in Philadelphia and three of four did in Los Angeles. In New York, high wind velocity is again important (both summers). Humidity is now low in Philadelphia (one of two summers). In Los Angeles, relative humidity and radiation are again negative (3 of 4 and 2 of 4 seasons respectively).

The regressions for winters combined and summers combined for the three cities (for total and 65+ mortality; tables 1 -3) yielded somewhat lower correlations than the season-specific regression. The New York total winters' correlation was higher, as was the Los Angeles' total summers. Some changes occurred in these Los Angeles results in the meteorological variables. More importantly, in Philadelphia, particulate matter was significant for both sets of seasons. The similarities between the cities were still present: Temperature and pollution were still consistently important.

Similarities in the relationship between mortality and air pollution-weather are indicated: temperature is almost always important, as is the presence of some pollution. Wind velocity and relative humidity are sometimes important, but their interactions with each other and the interactions of both with temperature (34) during different environmental conditions complicate the matter. One can say that they are important, but not as significant as tempera-

ture. The comparisons between cities also differentiated between the two types of atmospheres, reducing and photo-oxidant: In Los Angeles, the major pollutant to emerge was NO_x, a photo-oxidant product; SO_2 was also important. In New York the most important pollutants were SO_2 and particulate matter, which are reducing-type pollutants.

When lags were explored, the multiple correlation coefficients sometimes increased with one day lags but not consistently. They occasionally also increased with three day lags. They definitely decreased after three days of lag and up to 15 days of lag. It does fit with some of the evidence on mortality, as discussed by McCarroll and Bradley (43). In other words, the effect on mortality is almost immediate. The multiple correlation coefficients with one day lags are: about 0.4 for the total period for New York, about 0.18 for the total period for Philadelphia, and about 0.36 for the total period for Los Angeles. By season, they ranged from 0.47 to 0.69 for New York, from 0.15 to 0.37 for Philadelphia, and from 0.27 to 0.60 for Los Angeles. These results indicate that one to three day lag periods are useful in analyzing the relationship.

When one day lags were used, a few changes from zero lag results occurred in the variables present in the regressions by season for the different cities (Tables 1-3). This is due, at least in part, to the nature of the interactions discussed above. In New York, particulate matter was more important, as was temperature, and wind velocity became much less important. It is hard to explain why SO_2 was important during zero lag and PM was during lag one. It might again be a function of the time relationship in their interaction. For Philadelphia, the relationships with lag one were similar to those found for New York. For Los Angeles with one day of lag, NO_x became much more important (and SO_2 much less), while temperature was more consistently important, and relative humidity had about the same importance as previously. There was again some consistency between seasons and also some between cities.

It is to be noted that some pollutants had negative coefficients in some of the regression results. This too is a function of the specific interactions between variables in the specific conditions of the environment. The independent variables in multivariate analyses operate as a set. There are conditions in which the pollutants will be lower than usual, either due to their decreased production, their removal by the weather conditions prevalent at that time, or their failure to stay in the atmosphere for these reasons. In such cases which prove to be significantly related to increased mortality, the pollutant components in the equation will be negative or will not show, depending on the circumstances (discussed above). A good example is the condition of stormy weather in winter (say in New York), where temperature will often be low, wind velocity will be high, there probably will be some precipitation, and barometric pressure will often be low, thus preventing anti-cyclonic inversion conditions and in fact removing the pollution from the air. In time, the interaction of variables will produce meteorological conditions in which one weather variable will have increasing effects on other weather variables, and possibly health, and it will then decrease in effect while the other weather variables may have increasing effect on health. It appears from this analysis that more data, both years of data and number of variables, is needed for Philadelphia.

Conclusions

Generally then, the relationships found for each city for each season involved both the pollutants and the meteorological variables. Some consistency was found between seasons within each city and between cities, but large differences also exist. Each season in each year must be considered as somewhat unique. Comparisons between cities showed some consistencies. Even when different, the cities showed similarities in the relationships for winters. The differences were somewhat magnified during the summers and were

in accord with their differences in environment. Only about 25% of the variability in mortality was explained by the regressions in New York; about 15% was explained in Los Angeles, and less in Philadelphia. This leaves a great deal of variability in mortality to be explained. This may be a function of the method used. It is also because pollution and weather are only two of the possible "causes" of mortality.

TABLE 1

Coefficients in the Step-Wise Multiple Regression Analysis* of Mortality on Stated Environmental Variables by Season Using Standard Score Variables.

New York: July 1962 - June 1965

Total period	SO^2	Particulate Matter	CO	NO	Temp.	Rel. HUm.	Wind Vel.	Radiation	Multiple R	R^2 (%) **
Total mortality	+	+.39			-.10		+.20		.380	14.4%
65+ Age mortality		+.37			-.13		+.16		.374	14.0%
Seasons (Total Mortality)										
Oct. 62 - Mar. '63	1.07	[illegible]		-.23	-.74	.29	.46		.834	69.6%
Apr.- Sept. '63	-.33	.34			.22		+		.398	15.8%
Oct. '63- Mar '64		.49					.31		.413	17.0%
Apr. - Sept. ' 64	.35		.26				.18		.520	27%
Oct. '64 - Mar. '65				+	-.27			-.29	.378	14.3%
				One Day Lag						
Total period										
Total mortality		.34			-.21	-.11			.409	16.7%
65+ Age mortality		.34			-.21	-.11			.413	17.0%
Seasons (Total Mortality)										
Oct'62 - Mar '63		.45			-.53				.687	47.2%
Apr. - Sept. '63	-.34	.53			.31	-.26			.615	37.8%
Oct. '63-Mar '64		.34			-.33				.473	22.4%
Apr. - Sept. ' 64	.25	.28				-.22			.501	25.1%
Oct. '64 - Mar. '65	.20				-.33			-.25	.477	22.8%

*Criteria $F_{out} = 2.75$, all multiple R are significant, $P < .01$

**Square of multiple correlation coefficient in % = amount of variability in mortality "explained"

+ t(.05-.10) coeff. <.1

TABLE 8

Coefficients in the Step-Wise Multiple Regression Analysis* of Mortality on Stated Environmental Variables by Season Using Standard Score Variables.

Philadelphia: 1963 - 1964

	Particulate Matter TSP	COH	Temp.	Rel Hum.	Multiple R	R^2(%)**
Total period						
Total mortality		.16			.165	2.7%
65+ Age mortality		.14			.137	1.9%
Seasons (Total Mortality)						
Apr. - Sept. '63			.28	-.18	.272	7.4%
Oct. '63 - Mar '64			-.33	.21	.355	12.6%
Apr. - Sept. '64	.19				.193	3.7%
			One Day Lag			
Total Period						
Total mortality	.16		-.12		.184	3.4%
65+ Age mortality	.16		-.09		.170	2.9%
Seasons (Total Mortality)						
Apr. - Sept '63			.16		.164	2.7%
Oct. '63 - Mar ' 64			-.37	.15	.372	13.8%
Apr. - Sept. '64	.15				.151	2.3%

*Criteria $F_{.05}$, $t_{.05}$, all multiple R are significant, P>.01

**Square of multiple correlation coeficient in % = amount of variability in mortality "explained"

TABLE 3

Coefficients in the Step-Wise Multiple Regression Analysis* of Mortality on Stated Environmental Variables by Season Using Standard Score VAriables.
Los Angeles: 1962 - 1965

	SO_2	CO	NO_X	Temp.	Rel Hum.	Wind Vel.	Radiation	Multiple R	R^2 (%) **
Total period									
Total mortality	.11	-.11	.29		-.10	.10	-.12	.337	11.4%
65+ Age mortality	.10	-.14	.40		†	.11	-.06	.344	11.8%
Seasons									
(Total Mortality)									
Apr. - Sept. '62			.18		-.28			.400	16%
Oct. ' 62 - Mar. '63	.24		.23			.30	-.28	.444	19.7%
Apr. - Sept. '63	.30			.22	-.32		-.16	.539	29.0%
Oct. '63 - Mar ' 64	.19			-.23	-.20			.371	13.8%
Apr. - Sept. '64				†	-.29		-.23	.265	7.0%
Oct. '64 - Mar '65			.31	-.20				.365	13.3%
Apr. - Sept. '65		-.23	.38					.282	8.0%

One Day Lag

Total period								
Total mortality	.07	.20	-.14	-.20	.06	-.07	.361	13.0%
65+ Age mortality	.08	.14	-.14	-.25		-.08	.354	12.5%
Seasons								
(Total Mortality)								
Apr. - Sept. '62		.27			†		.267	7.1%
Oct. ' 62 - Mar. '63		.17	-.18	-.28			.432	18.7%
Apr. - Sept. '63	.13	.29		-.32	.22	†	.601	36.1%
Oct. '63 - Mar '64			-.27	-.36		†	.435	18.9%
Apr. - Sept. '64	†	.28	-.22				.308	9.5%
Oct. '64 - Mar '65		-.29	-.20				.409	16.7%
Apr. - Sept. '65				-.31			.310	9.6%

*Criteria $F_{.05}$, $t_{.05}$, all multiple R are significant, $P<.01$
**Square of multiple correlation coefficient in % = amount of variability in mortality "explained"
† $t_{.05-.10}$, coeff. <.1
††No data available

Bibliography

1. Schrenk, H.H., Heimann, H., Clayton, G.D., Gafaffer, W.M., and Wexler, H. Air Pollution in Donora, Pa. Epidemiology of the unusual smog episode of October, 1948. Publ. Hlth. Bull. (Wash.). 1949, No. 306.

2. Bradley, W.H., Logan, W.P.D., and Martin, A.E. The London Fog of December 2nd-5th. Office of the Ministry of Health, 1957, Her Majesty's Stationary Office, pp. 156-165.

3. Firket, J. Fog along the Meuse Valley. Trans. Faraday Soc., 1936, 32: 1192-1197.

4. Breslow, L. Atmospheric Conditions of Cities in Clinical Cardiopulmonary Physiology. Grune & Stratton, Inc., 1960, pp. 926-935.

5. Ferris, B.G., Jr. Chronic Low-Level Air Pollution. Use of General Mortality, and Chronic Disease Morbidity and Mortality to Estimate Effect. Environ. Res., 1969, 2:79-87.

6. Daly, C. Air Pollution and Causes of Death. Brit. J. prev. soc. Med., 1959, 13:14-27.

7. Schiffman, R., and Landau, E. Use of Indexes of Air Pollution Potential in Mortality Studies. Presented at the Air Pollution Medical Program, Division of Special Health Services, Department of Health, Education, and Welfare, Washington, D.C.

8. Hickey, R.J., Schoff, E.P. Relationship Between Air Pollution and Certain Chronic Disease Death Rates. Arch. Environ. Hlth., 1967, 15:728-738.

9. Winkelstein, W., Jr., Kantor, S., Davis, E.W., Maneri, C.S. and Mosher, N.E. The Relationship of Air Pollution and Economic Status to Total Mortality and Selected Respiratory System Mortality in Men. I. Suspended Particulates. Arch. Environ. Hlth., 1967, 14:162-171.

10. Winkelstein, N., Jr., Kantor, S., Davis, E.W., Maneri, C.S. and Mosher, N.E. The Relationship of Air Pollution and Economic Status to Total Mortality and Selected Respiratory System Mortality in Men. II. Oxides in Sulfur. Arch. Environ. Hlth., 1968, 16:401-405.

11. Zeidberg, L.D., Horton, R.J. and Landau, E. The Nashville Air Pollution Study. V. Mortality from Diseases of the Respiratory System in Relation to Air Pollution. Arch. Environ. Hlth., 1967, 15:214-224.

12. Lepper, M.H., Shioura, N., Carnow, B., Andelman, S., and Lehrer, L. Respiratory Disease in an Urban Environment. Indust. Med., 1969, 38: 126-131.

13. Carnow, B. Mortality and Pollution. Presented at APHA, Houston, 1970 and AMA Air Pollution Medical Research Conference, 1970.

14. Rumford, J. Mortality Studies in Relation to Air Pollution. Amer. J. Publ. Hlth., 1961, 51:165-173.

15. Scott, J.A., Taylor, I., Gore, A.T., and Shaddick, C.W. Mortality in London in the Winter of 1962-63. Med. Officer, 1964, 111:327-330.

16. Boyd, J.T. Climate, Air Pollution, and Mortality. Brit. J. prev. soc. Med., 1960, 14:123-135.

17. Hechter, H.H. and Goldsmith, J.R. Air Pollution and Daily Mortality. Amer. J. Med. Sci., 1961, 66:581-588.

18. Martin, A.E., and Bradley, W.H. Mortality, Fog and Atmospheric Pollution. An Investigation during the Winter of 1958-59. Office of the Ministry of Health, 1960, Her Majesty's Stationary Office, pp. 56-72.

19. Greenburg, J., Field, F., Erhardt, C., Glasser, M. and Reed, J. Air Pollution, Influenza and Mortality in New York City. Arch. Environ. Hlth., 1967, 15:430.

20. Greenburg, L., Jacobs, M.B., Drolette, B.M., Field, F., and Braverman, M.M. Report of an Air Pollution Incident in New York City, November, 1953. Publ. Hlth. Reports, 1962, 77:7-16.

21. Glasser, M., and Greenburg, L. Air Pollution, Mortality, and Weather. Presented at the Annual Meeting of the American Public Health Association, Philadelphia, Nov. 11, 1969.

22. Hodgson, T.A. Short-term Effects of Air Pollution on Mortality in New York City. Env. Sci. & Techn., 1970, 4:589-97.

23. McCarroll, J.R. and Bradley, W. Excess Mortality as an Indicator of Health Effects of Air Pollution. Amer. J. Publ. Hlth., 1966, 56: 1933-1942.

24. Cassell, J.E., Wolter, D.W., Mountain, J.D., Diamond, J.R., Mountain, I.M., and McCarroll, J.R. II. Reconsiderations of Mortality as a Useful Index of the Relationship of Environmental Factors to Health. Amer. J. Publ. Hlth., 1968, 58:1653-1657.

25. Speizer, F.E. An Epidemiological Appraisal of the Effects of Ambient Air on Health: Particulates and Oxides of Sulfur. J. Air Pollut. Contr. Ass., 1969, 19:647-655.

26. Ipsen, J. Episodic Morbidity and Mortality in Relation to Air Pollution. Presented at the New York Academy of Sciences "Working Group on Epidemiology of Air Pollution and Human Disease", Dec. 1967.

27. Serfling, R.E. The Current Mortality Chart. Morb. and Mort. Wkly. Rep., 1965, 14:8.

28. Serfling, R.E. Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths. Publ. Hlth. Rep., 1963, 78:494-506.

29. Reinke, W.A. Biostatistical Aspects of the Air Pollution Health Problem. Arch. Environ. Hlth., 1967, 14:213-219.

30. Afifi, A.A. and Elashoff, R.M. Missing Observations in Multivariate Statistics. II. Point Estimation in Single Linear Regression. JASA, 1967, 62:10-29.

31. Anderson, T.W. An Introduction to Multivariate Statistical Analysis. New York:Wiley, 1958.

32. Dixon, W.J. (ed) Biomedical Computer Programs. Los Angeles:UCLA, 1965.

33. Toyama, T. and Adachi, S. An Interim Report of Study on Mortality and Air Pollution in Tokyo City. Given at project meeting of Japan-U.S. Science Cooperation Program, Jan., 1968.

34. Lowry, W.P. Weather and Life, 1968, Coravallis, Ore. OSU Press, Ch. 13-15.